Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s10803-025-06778-z

ORIGINAL ARTICLE

Sensitivity and Specificity Between the Screening Tool for Autism in Toddlers and Young Children (STAT) and Autism Diagnostic Observation Schedule (ADOS-2)

Jaclyn Gunderson[1] · Flora Howie[1] · Sunil Mehta[2] · Annie Redlin[1] · Andrea Huebner[3]

Accepted: 23 February 2025

Abstract
The current study evaluated the agreement of the Screening Tool for Autism in Toddlers & Young Children (STAT) in differential diagnosis of autism in an outpatient clinical population, compared to the more time and resource-intensive Autism Diagnostic Observation Schedule-Second Edition (ADOS-2). Sensitivity and specificity comparative analyses were completed on 44 patients (aged 24–36 months) who were administered both the STAT and the ADOS. Sensitivity and specificity were also calculated independently on patients that received a STAT (n = 102, 24–36 months) or ADOS-2 (n = 72, 24–36 months) and multidisciplinary clinical evaluation. Using clinical diagnosis as the measure of truth, 33 of the 44 received a clinical diagnosis of ASD. Agreement between the STAT and ADOS-2 was 90.9% (40/44; 95% CI 78.3–97.5%). The sensitivity of the STAT was 90.9% (30/33; 95% CI 75.7–98.1%) and the sensitivity of the ADOS was 100% (33/33, 95% CI 89.4–100%) in our sample. The specificity of the STAT was 90.9% (10/11; 95% CI 58.7–99.8%) and the specificity of the ADOS was 100% (11/11; 95% CI 75.1–100.0%). The STAT showed high sensitivity and moderate specificity in differentiating children with autism from those with other neurodevelopmental disorders in this outpatient clinic population. There was excellent agreement between the STAT and ADOS-2. The STAT may be an acceptable diagnostic tool to refine clinic models and reduce wait times for evaluation in toddlers who present with concerns for autism.

Keywords Autism · Diagnosis · Early identification · Sensitivity · Specificity

Sensitivity and Specificity Between the Screening Tool for Autism in Toddlers and Young Children (STAT) and Autism Diagnostic Observation Schedule (ADOS-2).

Autism spectrum disorder (ASD) affects 1 in 36 children (Maenner, 2023). The combination of high autism prevalence and dearth of qualified evaluators results in families routinely waiting 1 to 2 years for diagnostic assessments (Ahlers et al., 2019; Maenner, 2023). In fact, children with ASD receive their diagnosis later than children with other developmental disorders, despite the fact that they are often substantially younger when parents first report concerns about their development (Oswald et al., 2017). Delays to diagnosis are problematic because they lead to delays for early intervention (Oswald et al., 2017). Children that could benefit from intervention during critical brain plasticity periods are put on lengthy waitlists that can jeopardize developmental outcomes (Zwaigenbaum et al., 2015). Therefore, it is important to identify effective and efficient tools for early ASD identification.

Best practice ASD diagnostic evaluations often include an interdisciplinary team using standardized and validated measures. Yet, traditional evaluation measures contribute to the problem of long wait times due to considerable training requirements and extensive administration and scoring times resulting in fewer children seen due to longer evaluation times. For example, the Autism Diagnostic Observation Schedule (ADOS) is a semi-structured observation assessment considered to be a gold standard assessment tool for use in evaluating ASD (Lord et al., 2000). However, the ADOS takes 45 to 60 min to complete, plus additional time to score and interpret.

✉ Jaclyn Gunderson
Gunderson.Jaclyn@mayo.edu

[1] Department of Pediatric and Adolescent Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA

[2] PrairieCare, Rochester, MN, USA

[3] Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN, USA

Published online: 09 March 2025

The Screening Tool for Autism in Toddlers and Young Children (STAT) is a Level 2 screener for autism detection in children 24 to 36 months of age with extended norms down to age 14 months (Stone et al., 2004). Level 1 screeners are often implemented in primary care settings and include assessments such as the Ages and Stages Questionnaire®, Parent's Evaluation of Developmental Status®, and Child Development Inventory. In comparison to Level 1 screeners generally used to identify children at risk for developmental delays, Level 2 screeners are used to aid in distinguishing ASD risk from other developmental disorders once the child has already been flagged for concerns. Other Level 2 screeners include the Rapid Interactive Screening Test for Autism in Toddlers (RITA-T; Choureri & Wagner, 2015), and the Systematic Observation for Red Flags (SORF; Wetherby et al., 2004). The RITA-T is a 20-min play observation that probes joint attention, human agency, communication, social awareness, self-awareness and visual reception through 9 play-based activities. The SORF is rating scale of 22 behaviors associated with ASD that can be applied to videotaped observations of either the Behavior Sample of Communication Symbolic Behavior Scales (Wetherby & Prizant, 2002) or a 60-min home interaction (Delehanty & Wetherby, 2021). The SORF samples behaviors across social communication, social interactions and restricted and repetitive behaviors and interests. The STAT is like other observational Level 2 screeners in that a score is given for the lack of typical behavior and the presence of atypical behaviors during an interactive observation. However, the STAT offers a wider variety of play activities to assess key behaviors known to differentiate two-year-old children with ASD from those with other developmental delays. The STAT takes approximately 20 min to administer 12 play-based items that assess key social and communicative behaviors including imitation, play, requesting, and directing attention. Items are scored with pass/fail criteria that is averaged to obtain a domain score ranging from 0 to 1. The STAT can be administered by service providers in various roles (e.g., physicians, clinical social workers, nursing staff, research staff, counseling professionals, undergraduate trainees, graduate trainees). Yet, the STAT does not sample restricted and repetitive behaviors and is intended be used in conjunction with other relevant information gathered through clinical history or observation.

The original psychometric evaluation of the STAT as a Level 2 screener included a sample of 52 children ages 24 to 36 months-old, half of whom had a clinical diagnosis of autism (n = 26), and half who were diagnosed with developmental delay and/or language impairment (n = 26). Signal detection approach was used to determine the optimal cutoff score for identifying children at higher risk for autism, prioritizing sensitivity over specificity to minimize false negatives. Applying the cutoff score of 2 or higher to the validation sample the STAT had excellent sensitivity (0.92), specificity (0.85), positive predictive value (PPV; 0.86), and negative predictive value (NPV; 0.92) (Stone et al., 2004). Positive findings were replicated across larger and more diverse samples with good sensitivity (ranging 0.74—0.97) and specificity (ranging 0.70–0.93) (Chiang et al., 2013; Roberts et al., 2019; Rooney et al., 2022; Wu et al., 2020). In another study of 71 children downward extending the age range, the results again showed excellent sensitivity (0.93), specificity (0.85), PPV (0.86), and NPV (0.97) when using a higher cutoff score of 2.75 for children between 14 to 23 months (Stone et al., 2008). Beyond the work indicating the sensitivity and specificity of the STAT, Stone and colleagues observed high concurrent validity between the ADOS and STAT on categorizing ASD (kappa = 0.95) (Stone et al., 2004). Furthermore, Khowaja and colleagues reported significantly correlated total scores ($r = 0.51$, $p < 0.001$) and social affect scores ($r = 0.55$, $p < 0.001$) between the STAT and ADOS (Khowaja et al., 2018).

While previous evidence supports the use of the STAT as an ASD screener, it is possible that the STAT may be useful as a front-line diagnostic tool, especially for more clear-cut ASD cases. Previous research highlights the potential use of the STAT within Early Intervention settings to identify autism spectrum disorder (Tagavi et al., 2023). However, more work is needed to evaluate the diagnostic utility of the STAT and how it performs in comparison to current 'gold standard' ASD diagnostic tools. The current study evaluated the agreement of the STAT in differential diagnosis of an outpatient sub-specialty patient population, compared to the more time and resource-intensive Autism Diagnostic Observation Schedule-Second Edition (ADOS-2).

Methods

Developmental Behavioral Pediatrics Clinic patient log and EPIC software were used to identify 317 consecutive patients seen for diagnostic evaluation in a hospital outpatient clinic between September 1, 2019 and April 30, 2022. Of the 317 patient charts reviewed, 130 were included. Only patients between the ages of 24–36 months with a STAT or ADOS-2 administered during their clinical evaluation were included. Table 1 shows demographic variables for the participant sample. 44 patients received both the STAT and the ADOS-2, 58 separate patients received the STAT only, and 28 received the ADOS only. The STAT was administered by the nursing team and the ADOS-2 was administered separately by the speech/language pathology team. Scoring of the STAT and ADOS-2 was completed independently and based on the associated manual scoring algorithm. The ASD diagnosis was based on meeting DSM-5-TR criteria through a combination of clinical interview and history, behavioral

Table 1 Child demographics for those receiving either STAT, ADOS-2 or both assessments

	STAT + ADOS-2sample (n = 44)	STAT administration (n = 102)	ADOS-2 administration (n = 72)
Sex			
Female	11 (25%)	30 (29.4%)	18 (25%)
Male	33 (75%)	72 (70.5%)	54 (75%)
Age (months)			
Mean	30.59	29.72	30.5
Range	24–35	24–36	25–36
Ethnicity			
Hispanic	5 (11.3%)	12 (11.7%)	8 (11.1%)
Not Hispanic	39 (88.6%)	89 (87.3%)	63 (87.5%)
Unknown/choose not to disclose	0 (0%)	1 (.01%)	1 (3.5%)
Race			
American Indian/Alaskan Native	2 (4.5%)	4 (3.9%)	0 (0%)
Asian	4 (9.1%)	8 (7.8%)	6 (8.3%)
Black/African American	6 (13.6%)	17 (16.6%)	14 (19.4%)
White	30 (68.1%)	67 (65.6%)	45 (53.6%)
Other	0 (0%)	3 (2.9%)	0 (0%)
Unknown/choose not to disclose	2 (4.5%)	2 (2.9%)	5 (6.9%)

observation, standardized test results, and input from Speech-Language Pathologists, Psychiatrists, Psychologists, social workers, nurses, and Developmental Behavioral Pediatricians. While STAT and ADOS-2 administration was completed independently by different examiners, the results were used to inform the clinical diagnosis and therefore cannot be considered independent of the clinical diagnosis.

Measures

The Screening Tool for Autism in Toddlers and Young Children (STAT) was administered by nursing staff that completed the one-day STAT training tutorial and post-test provided through VU e-Innovations. STAT administration took place during outpatient office visits. The STAT testing was conducted in outpatient clinic rooms either on the floor or at a small table depending on the child's interest. Parents remained in the room during testing and were informed that the examiner would be trying to get the child to interact and that it is most helpful if the parent observed interactions and redirected the child back to the examiner if the child came to them. Administration of the STAT includes toys or objects being presented across several trials and the examiner scoring each trial based on the child's behavior. Turn taking play includes rolling a ball or toy car back and forth. Functional play is administered by providing a doll or stuffed animal with mealtime items or furniture and observing how the child manipulates the items (baby in the chair, feeding the teddy bear, etc.). Requesting behaviors are observed during bubble play and snack time. In both activities the examiner observes how the child requests more or asks for help getting the container open. Four different items are used to assess the child's behavior related to directing attention including the examiner inflating a balloon and letting it go, the examiner putting an animal puppet on his/her hand and writing notes, a bag of toys being given to the child and encouraging the child to look in the bag, and a noisemaker being activated outside of the child's view. Imitation items include modeling an action along with a designated vocalization and encouraging the child to do the same (shaking a rattle, rolling a car back and forth, drumming hands on the table, making an animal hop across the table). To calculate a child's summary score, the number of failures for each of the four sections (play, requesting, directing attention, and imitation were counted and scored based on the STAT scoring algorithm. At-risk scoring cut offs were based on the STAT manual and supported by previous psychometric evidence (Stone et al., 2004) with a score of 2 or higher resulting in an 'at-risk' classification.

The Autism Diagnostic Observation Schedule, Second Edition (ADOS-2; Lord et al., 2012) is a semi-structured, standardized measure assessing communication, social interaction, play/imagination, and restricted and/or repetitive behaviors resulting in a Social Affect (SA), Repetitive Restrictive Behavior (RRB), and total cut-off score. The ADOS-2 was administered by trained Speech-Language

Pathologists during a 60-min office visit. ADOS modules were administered based on the individuals' level of expressive language and chronological age. The algorithm for the toddler module has three risk ranges indicating the extent to which the child is demonstrating behaviors consistent with ASD. A score in either the range of 'mild-to-moderate concern' or 'moderate to severe concern' indicates the child showed behaviors consistent with ASD. Modules 1 and 2 again, use cut off scores to indicate classifications of either 'autism' or 'autism spectrum disorder'. An 'autism' ADOS-2 classification indicates the individuals scores are comparable to participants with autism who have similar levels of expressive language. An 'autism spectrum disorder' ADOS-2 classification indicates that the individual shows behavior consistent with autism but with less severity. The ADOS-2 includes play-based tasks such as 'bubble play', 'birthday party play', and 'using toy objects' intended to provide the examiner with information on social communication, play and stereotyped behavior. The ADOS-2 shows strong psychometric properties across several studies (Gotham et al., 2008, 2009; Molloy et al., 2011; Oosterling et al., 2010; Zander et al., 2016). The Toddler module, for children under 30 months, shows evidence of strong internal consistency ($\alpha = 0.90$ for SA; $\alpha = 0.5$ for RRB) and inter-rater reliability (kappa = 0.60) as well as sensitivity between 0.81–0.91 and specificity between 0.86–0.94 for verbal and non-verbal children (Luyster et al., 2009). The inter-rater reliability for items across modules 1 to 4 was 0.74–0.83. For module 1 sensitivity is reported between 0.95–0.97 and specificity between 0.50–0.94. For modules 2 sensitivity is between 0.91–0.98 and specificity is between 0.84–0.93.

Analysis

STAT and ADOS-2 sensitivity and specificity were first calculated on patients who received both a STAT and ADOS-2 along with a multidisciplinary clinical evaluation. Overall agreement between the STAT and the ADOS-2 was computed based on matched classification categorizations across all participants that completed both the STAT and the ADOS-2. Agreement was indicated as 'not at risk' on the STAT and 'non-spectrum' on the ADOS-2 module 1 or 2 or 'little-to-no concern' on the ADOS-2 toddler module or 'at risk' on the STAT and 'autism spectrum/autism' on ADOS-2 module 1 or 2 or 'mild-to-moderate concern/moderate-to-severe concern' on the ADOS-2 toddler module. Sensitivity and specificity were calculated using the multidisciplinary ASD clinical diagnosis as a measure of truth.

Sensitivity and specificity were also calculated independently on patients that received a STAT (n = 102, 44 + 58 additional) or ADOS-2 (n = 72, 44 + 28 additional) and clinical evaluation. Sensitivity was calculated as the proportion of children with a clinical ASD diagnosis correctly identified to be 'at risk' based on the STAT cut off score of 2 or higher. Specificity was calculated as the proportion of children without ASD who were correctly identified as 'not at risk' using the STAT. ADOS-2 sensitivity was calculated as the proportion of children with a clinical ASD diagnosis correctly identified to be in the range of 'mild-to-moderate concern,' 'moderate to severe concern,' 'autism spectrum' or 'autism.' ADOS-2 specificity was calculated as the proportion of children deemed not to have a clinical ASD diagnosis correctly identified as 'little-to-no-concern' or 'non-spectrum.'

Results

Agreement between the STAT and ADOS-2 (n = 44) was 90.9% (40/44; 95% CI 78.3–97.5%) based on matched classification ('not at risk' by the STAT and 'non-spectrum' by the ADOS-2 module 1 or 2 or 'no concern' by the ADOS toddler module, or 'at risk' by the STAT and 'autism spectrum disorder' or 'autism' by the ADOS-2 module 1 or 2 or 'mild-to-moderate concern' or 'moderate to severe concern' by the ADOS-2 toddler module) (Table 2). Using clinical diagnosis as the measure of truth, 33 of the 44 received a clinical diagnosis of ASD. The sensitivity of the STAT was 90.9% (30/33; 95% CI 75.7–98.1%) and the sensitivity of the ADOS was 100% (33/33, 95% CI 89.4–100%) in our sample (Table 3). Based on an exact McNemar's test the sensitivity estimates between the STAT and ADOS-2 are not significantly different from each other (# of discordant pairs, 3 vs 0, p = 0.25). The specificity of the STAT was 90.9% (10/11; 95% CI 58.7–99.8%) and the specificity of the ADOS was 100%

Table 2 Agreement between STAT and ADOS-2 among children that received both during a diagnostic evaluation

	ADOS classification			
	Non-spectrum/ No concern	ASD/Mild-to-moderate concern	Autism/Moderate-to-severe concern	ADOS not administered
STAT: Not at risk	**10**	1	2	13
STAT: At risk	1	**4**	**26**	45
STAT not administered	13	3	12	

Note: Bold numbers indicate cases in agreement

Table 3 Sensitivity of the STAT and ADOS-2 (n = 44) based on clinical diagnosis as measure of truth

	ADOS classification			STAT sensitivity	ADOS sensitivity
	Non-spectrum/No concern	ASD/Mild-to-moderate concern	Autism/Moderate-to-severe concern		
STAT: Not at risk	0	1	2	90.9% (95% CI: %)	100% (95% CI: 89.4–100.0%)
STAT: At risk	0	4	26		

Table 4 Specificity of the STAT and ADOS-2 (n = 44) based on clinical diagnosis

	ADOS classification			STAT Specificity	ADOS Specificity
	Non-spectrum/No concern	ASD/Mild-to-moderate concern	Autism/Moderate-to-severe concern		
STAT: Not at risk	10	0	0	90.9% (95% CI: 58.7–99.8%)	100% (95% CI: 75.1–100.0%)
STAT: At risk	1	0	0		

(11/11; 95% CI 75.1–100.0%). See Table 4. Again, based on an exact McNemar's test the two specificity estimates are not significantly different from each other (# of discordant pairs, 1 vs 0, p = 1.00).

From September 2019 to April 2022, 58 additional patients were evaluated for a clinical diagnosis of ASD and received the STAT, but no ADOS-2. These evaluations typically incorporated the Childhood Autism Rating Scale (CARS) to inform clinical diagnosis. Another 28 patients between the ages of 24–36 months received the ADOS-2 but did not receive the STAT during their clinical evaluation for ASD. We calculated agreement, sensitivity and specificity of all cases that received the STAT (n = 102) or the ADOS-2 (n = 72) compared to clinical diagnosis as the measure of truth. The agreement between the STAT and clinical diagnosis was 89.2% (91/102; 95% CI 81.5–94.5%). The sensitivity of the STAT in this group of patients was 85.3% (71/76; 95% CI 85.3–97.8%) based on the matched classification of at-risk and a clinical diagnosis of ASD. The specificity of the STAT was 76.9% (20/26; 95% CI 56.3–91.0%). Of the six cases where the STAT rated the child 'at risk' and the clinical diagnosis did not include autism, four children received diagnoses related to speech and language delays, one received a mixed developmental delay diagnosis, and one was noted to have sensory and behavior problem challenges. The agreement between the ADOS-2 and clinical diagnosis was 95.8% (69/72; 95% CI 88.3–99.1%). The sensitivity was 94.2% (49/52; 95% CI 84.1–98.8%) and specificity was 100% (20/20; 95% CI 83.1–100%) for the ADOS-2 in this sample.

Discussion

The current study evaluated the utility of the STAT in differential diagnosis of ASD based on its agreement with the ADOS-2 and clinical ASD diagnosis determination.

Table 5 Agreement between the STAT and clinical diagnosis and ADOS-2 and clinical diagnosis

	Clinical diagnosis for autism		Agreement
	NO	YES	
STAT (n = 102)			
Not at risk	**20**	5	
At risk	6	**71**	89.2% (95% CI 81.5–94.5%)
ADOS-2 (n = 72)			
Non-spectrum	**20**	3	95.8% (95% CI 88.3–99.1%)
ASD	0	**9**	
Autism	0	**40**	

Note: Bold numbers indicate cases in agreement

This study also examined and compared the sensitivity and specificity of the STAT and the ADOS-2 in identifying ASD. There was excellent agreement between the STAT and ADOS-2 (90.9%) as well as the STAT and clinical diagnosis (89.2%) (See Table 5). In this outpatient clinic population, the STAT showed high sensitivity (85.3%−90.9%) and reasonable specificity (76.9%−90.9%) in differentiating children with ASD from those with other neurodevelopmental disorders. Sensitivity and specificity were not significantly different in the sample of children that received both the STAT and the ADOS-2 assessments. Results from this study suggest the STAT may be an acceptable diagnostic tool in children 24 to 36 months of age who present with concerns for ASD.

Our results add to previous reports of strong sensitivity and specificity of the STAT (Chiang et al., 2013; Khowaja et al., 2018; Stone et al., 2004, 2008; Wu et al., 2020, 2021). The STAT has a relatively short administration time, which could aid in more targeted yet still comprehensive evaluation of patients with concerns for ASD. Utilizing the STAT could

help decrease the wait time bottlenecks patients and their families experience when seeking out evaluation for possible ASD. Using the STAT in conjunction with clinical judgment of initial symptom presentation risk may decrease the likelihood of false positives and false negatives associated with using the STAT alone. The more time and resource-intensive ADOS-2 might be reserved for children with more complex or unclear autism presentations. Barbaresi and colleagues (Barbaresi et al., 2022) used a prospective diagnostic study, to evaluate clinical diagnoses of ASD by Developmental Behavioral Pediatricians (DBPs) with versus without the ADOS. DBP's evaluations were consistent with ADOS referenced evaluations in 90.0% of cases, suggesting that the ADOS is generally not required for diagnosis of ASD in young children by DBPs. DBPs and other qualified diagnosticians such as Psychologists can and should use clinical judgment to identify children for whom ADOS may be needed. Taken together, our findings suggest that the STAT can provide more objective data as an alternative to the ADOS, especially in more clear-cut presentations of ASD.

Limited evidence exists comparing agreement of other level 2 screeners against the ADOS-2. To date, the RITA and SORF are primarily used as an intermediate step before a full diagnostic evaluation with hopes to use information gleaned to speed the diagnostic process (Choueiri et al., 2021, 2023). Yet, Kong et al., 2021 reported significant correlations between the RITA-T and the ADOS-2 Total Score ($r = 0.61$, $p < 0.001$) in a sample of 35 children ages 18 months to 7 years. Overall diagnostic agreement between the RITA-T and the ADOS-2 was 80% (28/35). Using the ADOS-2 as the measure of truth, the RITA-T lacked sensitivity, with 20.6% false negatives (6/29). Of note, all 6 false negatives occurred in the older age group of 37–84-month-old children.

While the results suggesting the utility of the STAT in identifying ASD are promising, it is important to consider alternative scoring methods that may help increase diagnostic accuracy. For example, in a study of 109 children aged 16 to 30 months, when the scoring algorithm was updated to an alternative seven-item scoring equally weighting 7 of the 12 items that were most predictive of an ASD diagnosis, sensitivity was 0.78 (Khowaja et al., 2018). Using a sample of 171 children aged 24 to 36 months, Roberts and colleagues compared the performance of the STAT based on different scoring methods including, 1.) the original single cut off score, 2.) threshold based on the 7 most predictive items, and 3.) logistic regression model of all items scored and weighted depending on how well the item predicted ASD diagnosis (Roberts et al., 2019). Results of this study suggested that using a two-threshold, logistic regression model provided the most diagnostic accuracy. This data suggests that probability-based estimates with two thresholds may help provide clearer information related to cases that have a high or low probability of meeting ASD criteria. Using a two-threshold scoring model may enable general pediatricians to make autism diagnoses in more clear-cut cases, which may greatly reduce wait times for specialty evaluations. Our study results showing high sensitivity of the STAT converge with the idea that the STAT may be useful in identifying kids with high symptoms, yet the moderate specificity suggests that a two-threshold model might better support those with scores in the middle that need to go on for more specialty assessment.

Shortages of healthcare providers such as DBPs or Psychologists with specialized training in ASD diagnosis result in the need for effective and efficient alternative diagnostic evaluation models. In one study, community based primary care providers (PCPs) trained to use the STAT improved timeliness of diagnoses by an estimated two to six months and enhanced local access to care by increasing the number of primary care providers accepting referrals for autism diagnostic evaluations (Mazurek et al., 2019). PCP's from Missouri participated in the ECHO Autism STAT project to help establish and expand diagnostic capabilities to underserved regions of the state. Results from this study showed that the PCP participants achieved reliable administration of the STAT (89%) and increased their capacity for autism care and diagnosis (Mazurek et al., 2019). Using a systematic assessment tool such as the STAT may also help satisfy the stipulations of insurance companies and government payors who often require standardized measures to provide coverage for services. Future research should continue to evaluate the usefulness of the STAT in diagnostic evaluations, especially within community care practices that have the potential to increase access to early autism diagnosis in underserved populations.

The implications of our findings should be considered within the context of the limitations of this research. Our analyses are based on a convenience sample with assessment attainment being at the discretion of the diagnosing physician. The autism diagnosis was determined by a multidisciplinary team according to DSM-5-TR criteria but individuals' scores on the STAT and the ADOS-2 were considered in this process. Therefore, these determinations are difficult to interpret independently. A prospective, blinded case-controlled study would be preferred for future evaluations. However, in our study, like others evaluating alternative assessment options in evaluating ASD, there is knowledge to be extracted. Hine et al., 2020 evaluated diagnostic wait times using a streamlined process including medical/developmental history, ASD-specific questions on sensory concerns, behavioral issues, and STAT administration. For subjects requiring further diagnostic clarity, the ADOS-2, cognitive/adaptive testing and other standardized rating scales were added. Hine and colleagues found agreement between the STAT and ADOS in seven of the nine cases requiring

diagnostic clarity. It is of note that although the case number was small, the agreement between the STAT and ADOS-2 was good (78%), as was our finding in the current study (90.9%). Although good, it is possible Hein's agreement was lower as it selected the more challenging subjects requiring diagnostic clarity. The ADOS-2 and STAT as utilized in our study and the study by Hine were not diagnostic in isolation but synthesized with multiple other variables in keeping with best practice. This included a clinician trained to diagnose autism informed by the structure of DSM-5 diagnostic criteria (Hyman et al., 2020). This suggests that we can glean clinically meaningful information from our study that seeks to explore a less time and resource-intensive alternative for autism evaluations.

It is important to note that the STAT items do not consider all diagnostic criteria for an ASD diagnosis based on the DSM-5-TR. While the STAT may be useful in providing a structured observation of common ASD symptoms in children up to three years of age, it should only be used in conjunction with other relevant clinical information (i.e., clinical history, repetitive and restricted behavior patterns).

Conclusion

Results of this study suggest that the STAT is a useful assessment of ASD that can be administered in less time and with less training than other standardized observation assessments. Using the STAT in cases with providers indicating high diagnostic certainty may reduce the wait time for specialty ASD diagnostic evaluations by completing clear cut evaluations in a more efficient manner. Evaluations using more time intensive assessments such as the ADOS-2 may be reserved for cases that are more ambiguous or more complex. These findings highlight the importance of considering alternative tools that may streamline diagnostic evaluations.

Author Contributions All authors contributed to the study conception and design. Material preparation, and data collection were performed by Annie Redlin, Flora Howie. The first draft of the manuscript was written by Jaclyn Gunderson. Supervision of this project was provided by Andrea Huebner. All authors provided writing review and edits. All authors read and approved the final manuscript.

Funding Not applicable.

Data Availability Data generated for this article are available from the corresponding author on reasonable request.

Materials Availability Data generated for this article are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no competing interests.

Ethical Approval Ethics approval was obtained from Mayo Clincs internal review boards.

Consent to Participate Ethics approval was obtained from Mayo Clincs internal review boards.

Consent for Publication Not applicable.

References

Ahlers, K., Gabrielsen, T. P., Ellzey, A., Brady, A., Litchford, A., Fox, J., Nguyen, Q.-T., & Carbone, P. S. (2019). A pilot project using pediatricians as initial diagnosticians in multidisciplinary autism evaluations for young children. *Journal of Developmental & Behavioral Pediatrics, 40*(1), 1. https://doi.org/10.1097/DBP.0000000000000621

Barbaresi, W., Cacia, J., Friedman, S., Fussell, J., Hansen, R., Hofer, J., Roizen, N., Stein, R. E. K., Vanderbilt, D., & Sideridis, G. (2022). Clinician diagnostic certainty and the role of the autism diagnostic observation schedule in autism spectrum disorder diagnosis in young children. *JAMA Pediatrics, 176*(12), 1233–1241. https://doi.org/10.1001/jamapediatrics.2022.3605

Chiang, C.-H., Wu, C.-C., Hou, Y.-M., Chu, C.-L., Liu, J.-H., & Soong, W.-T. (2013). Development of T-STAT for early autism screening. *Journal of Autism and Developmental Disorders, 43*(5), 1028–1037. https://doi.org/10.1007/s10803-012-1643-4

Choueiri, R., Garrison, W. T., Tokatli, V., Daneshvar, N., Belgrad, J., Zhu, G., & Zhang, B. (2023). The RITA-T (rapid interactive screening test for autism in toddlers) community model to improve access and early identification of autism in young children. *Child Neurology Open*. https://doi.org/10.1177/2329048X231203817

Choueiri, R., Lindenbaum, A., Ravi, M., Robsky, W., Flahive, J., & Garrison, W. (2021). Improving early identification and access to diagnosis of autism spectrum disorder in toddlers in a culturally diverse community with the rapid interactive screening test for autism in toddlers. *Journal of Autism and Developmental Disorders, 51*(11), 3937–3945. https://doi.org/10.1007/s10803-020-04851-3

Choueiri, R., & Wagner, S. (2015). A new interactive screening test for autism spectrum disorders in toddlers. *Journal of Pediatrics, 167*(2), 460–466. https://doi.org/10.1016/j.jpeds.2015.05.029

Delehanty, A. D., & Wetherby, A. M. (2021). Rate of communicative gestures and developmental outcomes in toddlers with and without autism spectrum disorder during a home observation. *American Journal of Speech-Language Pathology, 30*(2), 649–662.

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(5), 693–705. https://doi.org/10.1007/s10803-008-0674-3

Gotham, K., Risi, S., Dawson, G., Tager-flusberg, H., Joseph, R., Carter, A., Hepburn, S., McMAHON, W., Rodier, P., Hyman, S. L., Sigman, M., Rogers, S., Landa, R., Spence, M. A., Osann, K., Flodman, P., Volkmar, F., Hollander, E., Buxbaum, J., … Lord, C. (2008). A Replication of the autism diagnostic observation schedule (ados) revised algorithms. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(6), 642–651. https://doi.org/10.1097/CHI.0b013e31816bffb7

Hine, J. F., Allin, J., Allman, A., Black, M., Browning, B., Ramsey, B., Swanson, A., Warren, Z. E., Zawoyski, A., & Allen, W. (2020). Increasing access to autism spectrum disorder diagnostic consultation in rural and underserved communities: Streamlined evaluation within primary care. *Journal of Developmental & Behavioral Pediatrics, 41*(1), 16–22.

Hyman, S. L., Levy, S. E., Myers, S. M., Kuo, D. Z., Apkon, S., Davidson, L. F., Ellerbeck, K. A., Foster, J. E. A., Noritz, G. H., Leppert, M. O. 'C., Saunders, B. S., Stille, C., Yin, L., Weitzman, C. C., Childers, D. O., Levine, J. M., Peralta-Carcelen, A. M., Poon, J. K., Smith, P. J., … Bridgemohan, C. (2020). Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics*. https://doi.org/10.1542/peds.2019-3447

Khowaja, M., Robins, D. L., & Adamson, L. B. (2018). Utilizing two-tiered screening for early detection of autism spectrum disorder. *Autism, 22*(7), 881–890. https://doi.org/10.1177/1362361317712649

Kong, X. J., Sherman, H. T., Tian, R., Koh, M., Liu, S., Li, A. C., & Stone, W. S. (2021). Validation of rapid interactive screening test for autism in toddlers using autism diagnostic observation schedule™ second edition in children at high-risk for autism spectrum disorder. Frontiers in Psychiatry, 12, 737890. https://doi.org/10.3389/fpsyt.2021.737890

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule—generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205–223. https://doi.org/10.1023/A:1005592401947

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism diagnostic observation schedule, second edition (ADOS-2)*. Western Psychological Services.

Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., Bishop, S., Esler, A., Hus, V., Oti, R., Richler, J., Risi, S., & Lord, C. (2009). The Autism diagnostic observation schedule—toddler module: A new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(9), 1305–1320. https://doi.org/10.1007/s10803-009-0746-z

Maenner, M. J. (2023). Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring Network, 11 Sites, United States, 2020. *MMWR. Surveillance Summaries*. https://doi.org/10.15585/mmwr.ss7202a1

Mazurek, M. O., Curran, A., Burnette, C., & Sohl, K. (2019). ECHO autism STAT: Accelerating early access to autism diagnosis. *Journal of Autism and Developmental Disorders, 49*(1), 127–137. https://doi.org/10.1007/s10803-018-3696-5

Molloy, C. A., Murray, D. S., Akers, R., Mitchell, T., & Manning-Courtney, P. (2011). Use of the autism diagnostic observation schedule (ADOS) in a clinical setting. *Autism, 15*(2), 143–162. https://doi.org/10.1177/1362361310379241

Oosterling, I., Roos, S., de Bildt, A., Rommelse, N., de Jonge, M., Visser, J., Lappenschaar, M., Swinkels, S., van der Gaag, R. J., & Buitelaar, J. (2010). Improved diagnostic validity of the ADOS revised algorithms: A replication study in an independent sample. *Journal of Autism and Developmental Disorders, 40*(6), 689–703. https://doi.org/10.1007/s10803-009-0915-0

Oswald, D. P., Haworth, S. M., Mackenzie, B. K., & Willis, J. H. (2017). Parental report of the diagnostic process and outcome: ASD compared with other developmental disabilities. *Focus on Autism and Other Developmental Disabilities, 32*(2), 152–160. https://doi.org/10.1177/1088357615587500

Roberts, M. Y., Stern, Y., Hampton, L. H., Grauzer, J. M., Miller, A., Levin, A., Kornfeld, B., Davis, M. M., Kaat, A., & Estabrook, R. (2019). Beyond pass-fail: Examining the potential utility of two thresholds in the autism screening process. *Autism Research, 12*(1), 112–122. https://doi.org/10.1002/aur.2045

Rooney, T., Stern, Y. S., Hampton, L. H., Grauzer, J., Hobson, A., Levin, A., Jones, M. K., Kaat, A. J., & Roberts, M. Y. (2022). Screening for autism in 2-year-old children: The application of the systematic observation of red flags to the screening tool for autism in toddlers and young children. *American Journal of Speech-Language Pathology, 31*(6), 2759–2769. https://doi.org/10.1044/2022_AJSLP-22-00132

Stone, W. L., Coonrod, E. E., Turner, L. M., & Pozdol, S. L. (2004). Psychometric properties of the STAT for early autism screening. *Journal of Autism and Developmental Disorders, 34*(6), 691–701. https://doi.org/10.1007/s10803-004-5289-8

Stone, W. L., McMahon, C. R., & Henderson, L. M. (2008). Use of the Screening tool for autism in two-year-olds (STAT) for children under 24 months: An exploratory study. *Autism, 12*(5), 557–573. https://doi.org/10.1177/1362361308096403

Tagavi, D. M., Dick, C. C., Attar, S. M., Ibanez, L. V., & Stone, W. L. (2023). The implementation of the screening tool for autism in toddlers in part C early intervention programs: An 18-month follow-up. *Autism, 27*(1), 173–187. https://doi.org/10.1177/13623613221086329

Wetherby, A., & Prizant, B. (2002). *Communication and symbolic behavior scales developmental profile -first* (normed). Paul H.Brookes.

Wetherby, A. M., Woods, J., Allen, L., Cleary, J., Dickinson, H., & Lord, C. (2004). Early indicators of autism spectrum disorders in the second year of life. *Journal of Autism and Developmental Disorders, 34*, 473–493.

Wu, C.-C., Chiang, C.-H., Chu, C.-L., Iao, L.-S., & Hou, Y.-M. (2021). T-STAT for detecting autism spectrum disorder in toddlers aged 18–24 months. *Autism, 25*(4), 911–920. https://doi.org/10.1177/1362361320972300

Wu, C.-C., Chu, C.-L., Stewart, L., Chiang, C.-H., Hou, Y.-M., & Liu, J.-H. (2020). The utility of the screening tool for autism in 2-Year-olds in detecting autism in Taiwanese toddlers who are less than 24 months of age: A longitudinal study. *Journal of Autism and Developmental Disorders, 50*(4), 1172–1181. https://doi.org/10.1007/s10803-019-04350-0

Zander, E., Willfors, C., Berggren, S., Choque-Olsson, N., Coco, C., Elmund, A., Moretti, Å. H., Holm, A., Jifält, I., Kosieradzki, R., Linder, J., Nordin, V., Olafsdottir, K., Poltrago, L., & Bölte, S. (2016). The objectivity of the autism diagnostic observation schedule (ADOS) in naturalistic clinical settings. *European Child & Adolescent Psychiatry, 25*(7), 769–780. https://doi.org/10.1007/s00787-015-0793-2

Zwaigenbaum, L., Bauman, M. L., Choueiri, R., Kasari, C., Carter, A., Granpeesheh, D., Mailloux, Z., Smith Roley, S., Wagner, S., Fein, D., Pierce, K., Buie, T., Davis, P. A., Newschaffer, C., Robins, D., Wetherby, A., Stone, W. L., Yirmiya, N., Estes, A., & Hansen, R. L. (2015). Early Intervention for children with autism spectrum disorder under 3 years of age: Recommendations for practice and research. *Pediatrics, 136*, S60–S81. https://doi.org/10.1542/peds.2014-3667E